



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

# UIT-T

SECTOR DE NORMALIZACIÓN  
DE LAS TELECOMUNICACIONES  
DE LA UIT

# G.711

**Apéndice II**  
(02/2000)

SERIE G: SISTEMAS Y MEDIOS DE TRANSMISIÓN,  
SISTEMAS Y REDES DIGITALES

Sistemas de transmisión digital – Equipos terminales –  
Codificación de señales analógicas mediante modulación  
por impulsos codificados (MIC)

---

Modulación por impulsos codificados (MIC) de  
frecuencias vocales

**Apéndice II: Definición de la cabida útil de  
ruido de confort para utilización según la  
Recomendación UIT-T G.711 en los sistemas  
de comunicaciones multimedios por paquetes**

Recomendación UIT-T G.711 – Apéndice II

(Anteriormente Recomendación del CCITT)

---

RECOMENDACIONES UIT-T DE LA SERIE G  
**SISTEMAS Y MEDIOS DE TRANSMISIÓN, SISTEMAS Y REDES DIGITALES**

|   |                    |
|---|--------------------|
| CONEXIONES Y CIRCUITOS TELEFÓNICOS INTERNACIONALES  | G.100–G.199        |
| <b>SISTEMAS INTERNACIONALES ANALÓGICOS DE PORTADORAS</b>  |                    |
| CARACTERÍSTICAS GENERALES COMUNES A TODOS LOS SISTEMAS ANALÓGICOS DE PORTADORAS   | G.200–G.299        |
| CARACTERÍSTICAS INDIVIDUALES DE LOS SISTEMAS TELEFÓNICOS INTERNACIONALES DE PORTADORAS EN LÍNEAS METÁLICAS  | G.300–G.399        |
| CARACTERÍSTICAS GENERALES DE LOS SISTEMAS TELEFÓNICOS INTERNACIONALES EN RADIOENLACES O POR SATÉLITE E INTERCONEXIÓN CON LOS SISTEMAS EN LÍNEAS METÁLICAS | G.400–G.449        |
| COORDINACIÓN DE LA RADIOTELEFONÍA Y LA TELEFONÍA EN LÍNEA   | G.450–G.499        |
| <b>EQUIPOS DE PRUEBAS</b>   |                    |
| <b>CARACTERÍSTICAS DE LOS MEDIOS DE TRANSMISIÓN</b>   | G.600–G.699        |
| <b>SISTEMAS DE TRANSMISIÓN DIGITAL</b>  |                    |
| EQUIPOS TERMINALES  | G.700–G.799        |
| Generalidades   | G.700–G.709        |
| <b>Codificación de señales analógicas mediante modulación por impulsos codificados (MIC)</b>  | <b>G.710–G.719</b> |
| Codificación de señales analógicas mediante métodos diferentes de la MIC  | G.720–G.729        |
| Características principales de los equipos multiplex primarios  | G.730–G.739        |
| Características principales de los equipos multiplex de segundo orden   | G.740–G.749        |
| Características principales de los equipos multiplex de orden superior  | G.750–G.759        |
| Características principales de los transcodificadores y de los equipos de multiplicación de circuitos digitales   | G.760–G.769        |
| Características de operación, administración y mantenimiento de los equipos de transmisión  | G.770–G.779        |
| Características principales de los equipos multiplex de la jerarquía digital síncrona   | G.780–G.789        |
| Otros equipos terminales  | G.790–G.799        |
| REDES DIGITALES   | G.800–G.899        |
| SECCIONES DIGITALES Y SISTEMAS DIGITALES DE LÍNEA   | G.900–G.999        |

Para más información, véase la Lista de Recomendaciones del UIT-T.

# RECOMENDACIÓN UIT-T G.711

## MODULACIÓN POR IMPULSOS CODIFICADOS (MIC) DE FRECUENCIAS VOCALES

### APÉNDICE II

#### **Definición de la cabida útil de ruido de confort para utilización según la Recomendación UIT-T G.711 en los sistemas de comunicaciones multimedios por paquetes**

#### **Resumen**

Este apéndice define un formato de cabida útil de ruido de confort (o tren de bits) para utilizar el códec de la Recomendación UIT-T G.711 en los sistemas de comunicaciones multimedios por paquetes.

El formato de la cabida útil está destinado a su utilización por sistemas por paquetes que tienen una gran tara de encabezamiento, donde la velocidad de transmisión de paquetes desempeña un papel significativo en la velocidad binaria global del sistema. En esta situación, el uso de algoritmos VAD/DTX/CNG puede reducir significativamente la velocidad de transmisión de paquetes y por tanto mejorar la eficacia de anchura de banda.

#### **Orígenes**

El apéndice II a la Recomendación UIT-T G.711 ha sido preparada por la Comisión de Estudio 16 (1997-2000) del UIT-T y fue aprobada por el procedimiento de la Resolución N.º 1 de la CMNT el 28 de febrero de 2000.

## PREFACIO

La UIT (Unión Internacional de Telecomunicaciones) es el organismo especializado de las Naciones Unidas en el campo de las telecomunicaciones. El UIT-T (Sector de Normalización de las Telecomunicaciones de la UIT) es un órgano permanente de la UIT. Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Conferencia Mundial de Normalización de las Telecomunicaciones (CMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución N.º 1 de la CMNT.

En ciertos sectores de la tecnología de la información que corresponden a la esfera de competencia del UIT-T, se preparan las normas necesarias en colaboración con la ISO y la CEI.

## NOTA

En esta Recomendación, la expresión "Administración" se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

## PROPIEDAD INTELECTUAL

La UIT señala a la atención la posibilidad de que la utilización o aplicación de la presente Recomendación suponga el empleo de un derecho de propiedad intelectual reivindicado. La UIT no adopta ninguna posición en cuanto a la demostración, validez o aplicabilidad de los derechos de propiedad intelectual reivindicados, ya sea por los miembros de la UIT o por terceros ajenos al proceso de elaboración de Recomendaciones.

En la fecha de aprobación de la presente Recomendación, la UIT no ha recibido notificación de propiedad intelectual, protegida por patente, que puede ser necesaria para aplicar esta Recomendación. Sin embargo, debe señalarse a los usuarios que puede que esta información no se encuentre totalmente actualizada al respecto, por lo que se les insta encarecidamente a consultar la base de datos sobre patentes de la TSB.

© UIT 2000

Es propiedad. Ninguna parte de esta publicación puede reproducirse o utilizarse, de ninguna forma o por ningún medio, sea éste electrónico o mecánico, de fotocopia o de microfilm, sin previa autorización escrita por parte de la UIT.

## ÍNDICE

### Página

|   |    |
|---|----|
| Apéndice II – Definición de la cabida útil de ruido de confort para utilización según la Recomendación UIT-T G.711 en los sistemas de comunicaciones multimedios por paquetes ..... | 1  |
| II.1 Alcance .....  | 1  |
| II.2 Definición de la cabida útil de ruido de confort .....   | 1  |
| II.2.1 Nivel de ruido .....   | 1  |
| II.2.2 Coeficientes de reflexión .....  | 2  |
| II.2.3 Empaquetamiento de la cabida útil.....   | 2  |
| II.3 Directrices de uso.....  | 3  |
| II.3.1 Factores que afectan a la calidad de funcionamiento del sistema.....   | 3  |
| II.3.2 Ilustración de las economías de anchura de banda en las aplicaciones a redes de paquetes .....   | 4  |
| II.4 Resultados de calidad de funcionamiento.....   | 5  |
| II.5 Ejemplo de solución.....   | 6  |
| II.5.1 Descripción del algoritmo.....   | 7  |
| II.5.2 Configuración probada .....  | 10 |



## Recomendación G.711

### MODULACIÓN POR IMPULSOS CODIFICADOS (MIC) DE FRECUENCIAS VOCALES

#### APÉNDICE II

#### Definición de la cabida útil de ruido de confort para utilización según la Recomendación UIT-T G.711 en los sistemas de comunicaciones multimedios por paquetes

(Ginebra, 2000)

##### II.1 Alcance

Este apéndice define un formato de cabida útil de ruido de confort (o tren de bits) para utilizar el códec de la Recomendación UIT-T G.711 en los sistemas de comunicaciones multimedios por paquetes. El formato de cabida útil es genérico y puede también utilizarse con otros códecs vocales sin capacidad de transmisión discontinua (DTX, *discontinuous transmission*) incorporada, como los de las Recomendaciones UIT-T G.726 [1], G.727 [2], G.728 [3] y G.722 [4]. El formato de la cabida útil proporciona una especificación de interoperabilidad mínima para la comunicación de parámetros de ruido de confort. El análisis y la síntesis de ruido de confort, así como los algoritmos de detección de actividad vocal (VAD, *voice activity detection*) y DTX no se especifican y siguen siendo específicos de la implementación. Sin embargo, se ha aprobado y se describe un ejemplo de solución. Utiliza el VAD y el DTX del anexo B/G.729 [5] y un algoritmo de generación de ruido de confort (CNG, *comfort noise generation*) que se proporciona como una información.

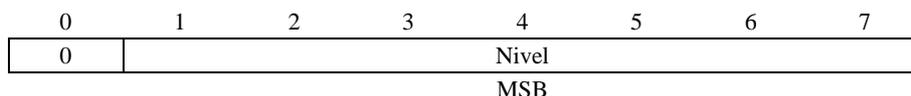
El formato de la cabida útil está destinado a su utilización por sistemas por paquetes que tienen una gran tara de encabezamiento, donde la velocidad de transmisión de paquetes desempeña un papel significativo en la velocidad binaria global del sistema. En esta situación, el uso de algoritmos VAD/DTX/CNG puede reducir significativamente la velocidad de transmisión de paquetes y por tanto mejorar la eficacia de anchura de banda.

##### II.2 Definición de la cabida útil de ruido de confort

La cabida útil de ruido de confort se compone de una descripción del nivel de ruido y de información espectral en forma de coeficientes de reflexión. El uso de información espectral es opcional y el orden del modelo todos polos no se especifica. El codificador puede determinar el orden del modelo apropiado basándose en consideraciones de calidad, complejidad, ruido ambiental previsto y anchura de banda de la señal. El orden del modelo no se transmite explícitamente, ya que puede obtenerse de la longitud de la cabida útil en el receptor. Por razones de complejidad o de otro tipo, el decodificador puede reducir el orden del modelo fijando a cero los coeficientes de reflexión de orden superior.

##### II.2.1 Nivel de ruido

El nivel de ruido se expresa en -dBov con valores de 0 a 127 que representan 0 a -127 dBov. dBov es el nivel relativo a la sobrecarga del sistema. El nivel de ruido se empaqueta con el bit más significativo (MSB, *most significant bit*) con el bit no utilizado siempre puesto a cero según la figura II.1.



**Figura II.1/G.711 – Empaquetamiento de bits de nivel de ruido**

### II.2.2 Coeficientes de reflexión

La información espectral se transmite mediante coeficientes de reflexión [6]. A partir del polinomio:

$$A(z) = 1 - \sum_{j=1}^M \alpha_j z^{-j}$$

obtenido por análisis de predicción lineal, puede obtenerse el conjunto de coeficientes de reflexión a partir del conjunto de coeficientes de LPC utilizando recursión hacia atrás de la forma:

$$k_i = -a_i^{(i)}$$

$$a_j^{(i-1)} = \frac{a_j^{(i)} + a_i^{(i)} a_{i-j}^{(i)}}{1 - k_i^2} \quad 1 \leq j \leq i-1$$

donde  $i$  baja de  $M$  a  $M-1$ , hasta 1 con la condición inicial:

$$a_j^{(M)} = \alpha_j \quad 1 \leq j \leq M$$

Obsérvese que esta formulación conduce a la solución a  $k_1$  expresada por:

$$k_1 = -\frac{r_i}{r_0}$$

donde  $r_i$  es el  $i$ -ésimo coeficiente de autocorrelación de la señal de entrada.

Cada coeficiente de reflexión puede tener valores entre  $-1$  y  $1$  y se cuantifica uniformemente mediante 8 bits. El valor cuantificado es representado por el índice  $N$  de 8 bits, donde  $N=0, \dots, 245$ , y el índice  $N=255$  se reserva para uso futuro. Cada índice  $N$  es empaquetado en un byte separado con el MSB primero. El valor cuantificado de cada coeficiente de reflexión puede obtenerse de su índice correspondiente por la fórmula:

$$\hat{k}_i(N_i) = \frac{258}{32768} \cdot (N_i - 127) \quad \text{para } N_i = 0, \dots, 254; -1 < \hat{k}_i(N_i) < 1$$

### II.2.3 Empaquetamiento de la cabida útil

El primer byte de la cabida útil debe contener el nivel de ruido como muestra la figura II.1. Los coeficientes de reflexión cuantificados se empaquetan en bytes diferentes por orden ascendente como en la figura II.2, donde  $M$  es el orden del modelo.

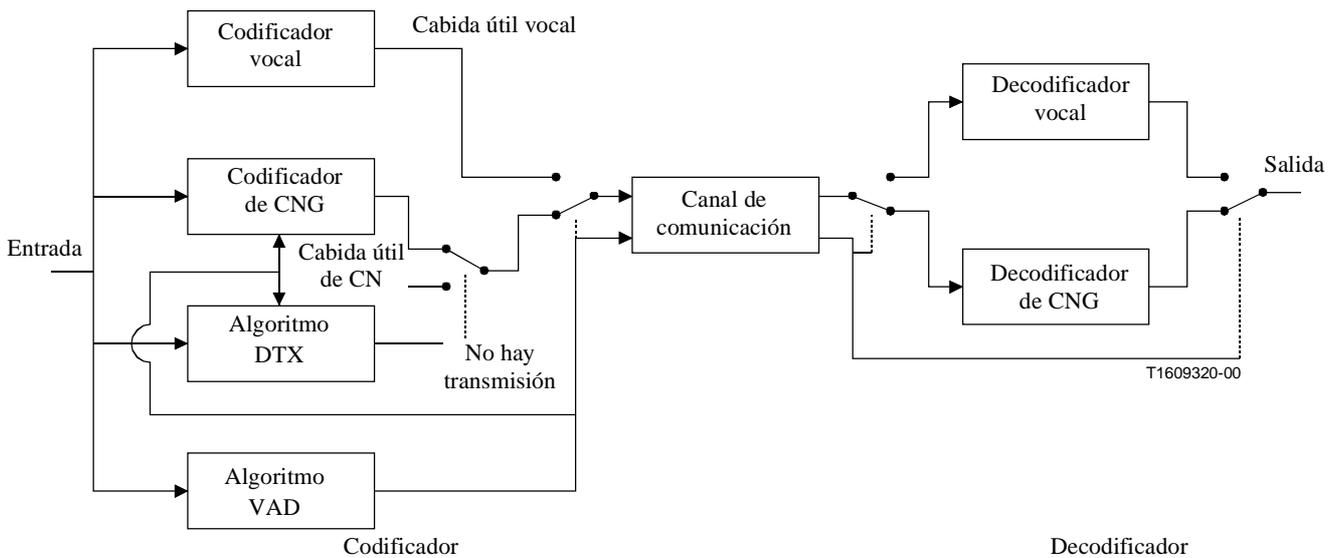


**Figura II.2/G.711 – Formato de empaquetamiento de la cabida útil de CN**

La longitud total de la cabida útil es  $M+1$  bytes. Obsérvese que un modelo de orden 0 (es decir, sin información de envolvente espectral) se reduce a transmitir únicamente el nivel de energía.

## II.3 Directrices de uso

La figura II.3 presenta el diagrama de bloques de un sistema de comunicación vocal con capacidades VAD/DTX/CNG. La misión del algoritmo VAD es discriminar entre segmentos de voz activa e inactiva en la señal de entrada. Durante los segmentos de voz inactiva, el papel del CNG es describir suficientemente el ruido ambiente, pero reduciendo al mínimo la velocidad de transmisión. Una trama de descriptor de inserción de silencio (SID, *silence insertion descriptor*) que contiene una descripción del ruido se empaqueta en la cabida útil de CN y se envía al receptor. El algoritmo DTX determina cuándo se transmite una trama SID. La trama SID puede enviarse periódicamente o sólo cuando hay un cambio significativo en la característica de ruido de fondo. El algoritmo CNG en el receptor utiliza la información del SID para actualizar su modelo de generación de ruido y producir luego una cantidad apropiada de ruido de confort.



**Figura II.3/G.711 – Sistema de comunicación vocal con DTX**

### II.3.1 Factores que afectan a la calidad de funcionamiento del sistema

La finalidad de los componentes VAD/DTX/CNG es reducir la velocidad de transmisión durante los periodos de señal vocal inactiva, pero manteniendo un nivel aceptable de calidad de salida. La calidad y la eficiencia son ambas afectadas por las prestaciones de cada uno de los componentes. Debe procurarse considerar conjuntamente las características de los algoritmos VAD, DTX y CNG, ya que de otro modo la calidad de funcionamiento obtenida por el sistema resultante podría ser deficiente.

#### II.3.1.1 VAD

El papel del algoritmo VAD es clasificar la señal de entrada en señal vocal activa y señal vocal inactiva o un ruido de fondo. La clasificación incorrecta de señal vocal inactiva como señal vocal activa tiene un efecto adverso en la eficiencia del sistema, al aumentar innecesariamente la velocidad de transmisión. En este caso, la calidad vocal no es afectada. Sin embargo, cuando la señal vocal activa se clasifica indebidamente como inactiva, se recorta la señal vocal y se degrada la calidad vocal. La mayoría de los algoritmos DTX emplean un periodo de retención cuando pasan de señal vocal activa a inactiva a fin de evitar recortar el extremo de cola de la señal vocal. Durante el periodo de retención, las tramas de señal vocal inactiva se reclasifican como señal vocal activa. El periodo de retención es también importante a fin de que el codificador de CNG obtenga una estimación exacta del ruido ambiente.

### II.3.1.2 DTX

El algoritmo DTX determina la frecuencia de la transmisión de tramas SID durante los periodos de señal vocal inactiva. Los esquemas DTX simples se actualizan periódicamente (por ejemplo, 5 Hz a 30 Hz). Los algoritmos DTX más complejos analizan la señal de entrada y transmiten sólo cuando se detecta un cambio significativo en el carácter del ruido ambiente [5].

### II.3.1.3 CNG

El papel del CNG es describir y reproducir el ruido ambiente. El ruido puede describirse adecuadamente por su energía y contenido espectral. A fin de evitar cambios bruscos en el carácter del ruido de confort, es importante promediar la estimación del parámetro en un periodo de tiempo. La cantidad de promediación apropiada depende del ruido ambiente, la calidad de funcionamiento y la retención del VAD, así como de la velocidad de actualización del DTX.

El orden del modelo utilizado es un factor en la exactitud de la estimación espectral. El orden óptimo es dependiente del ruido ambiente presente y de la anchura de banda de la señal. Es también importante adaptar el carácter espectral del ruido producido por el CNG con el del códec vocal. Por consiguiente, se sugiere que todo procesamiento previo de la señal de entrada antes del análisis dentro del codificador vocal se efectúe también dentro del codificador de ruido de confort.

## II.3.2 Ilustración de las economías de anchura de banda en las aplicaciones a redes de paquetes

El cuadro II.1 ilustra cómo el uso de transmisión discontinua en un sistema de comunicación por paquetes puede reducir significativamente la velocidad de transmisión y por ende mejorar la eficacia de anchura de banda. El ejemplo supone una tara de paquetes de 40 bytes, actividad vocal del 60% y una velocidad de actualización de DTX de 10 Hz.

**Cuadro II.1/G.711 – Economías de anchura de banda**

| Código | Velocidad binaria (bit/s) | Tamaño de paquete (ms) | Velocidad binaria IP (bit/s) | Cabida útil de CN de 1 byte        |               | Cabida útil de CN de 11 bytes      |               |
|--------|---------------------------|------------------------|------------------------------|------------------------------------|---------------|------------------------------------|---------------|
|        |                           |                        |                              | Velocidad binaria IP (prom. bit/s) | Economías (%) | Velocidad binaria IP (prom. bit/s) | Economías (%) |
| G.711  | 64 000                    | 5 ms                   | 128 000                      | 78 112                             | 39,0          | 78 432                             | 38,7          |
| G.711  | 64 000                    | 10 ms                  | 96 000                       | 58 912                             | 38,6          | 59 232                             | 38,3          |
| G.711  | 64 000                    | 20 ms                  | 80 000                       | 49 312                             | 38,4          | 49 632                             | 38,0          |
| G.726  | 32 000                    | 5 ms                   | 96 000                       | 58 912                             | 38,6          | 59 232                             | 38,3          |
| G.726  | 32 000                    | 10 ms                  | 64 000                       | 39 712                             | 38,0          | 40 032                             | 37,5          |
| G.726  | 32 000                    | 20 ms                  | 48 000                       | 30 112                             | 37,3          | 30 432                             | 36,6          |
| G.728  | 16 000                    | 5 ms                   | 80 000                       | 49 312                             | 38,4          | 49 632                             | 38,0          |
| G.728  | 16 000                    | 10 ms                  | 48 000                       | 30 112                             | 37,3          | 30 432                             | 36,6          |
| G.728  | 16 000                    | 20 ms                  | 32 000                       | 20 512                             | 35,9          | 20 832                             | 34,9          |

Por ejemplo, suponiendo un encabezamiento RTP/UDP/IP de 40 bytes, 60% de actividad vocal y una velocidad de actualización de DTX de 10 Hz, la velocidad binaria IP media con G.711 y una cabida útil CN de 11 bytes viene dada por:  $((64\ 000\ \text{bit/s}) + (40\ \text{bytes} \times 8\ \text{bit/byte} \times (1,0/0,005))) \times (0,6) + ((40+11)\ \text{bytes} \times 8\ \text{bit/byte} \times 10/\text{s}) \times (0,4) = 78\ 432\ \text{bit/s}$ .

## II.4 Resultados de calidad de funcionamiento

Se realizó una evaluación subjetiva de un ejemplo de implementación de CNG utilizando la cabida útil de CN. Se utilizó como método de evaluación el método de determinación de índices por categorías absolutas (ACR, *absolute category rating*) definido en la Recomendación UIT-T P.800. El material vocal utilizado en el experimento estaba compuesto por frases breves significativas y sencillas. El material de origen era IRS modificado filtrado (anexo D/UIT-T P.830) y organizado por pares. Cada par de frases duraba aproximadamente de 7 a 8 segundos, con un intervalo de tiempo entre frases de 1 segundo aproximadamente. En la evaluación se aplicaron condiciones de entrada calmadas y ruidosas, tales como ruidos de conversación confusa, de la calle, de oficina y de tráfico.

El códec vocal utilizado en el experimento era un códec G.711, procesado por el procedimiento de la figura II.4. En este experimento, la implementación sólo disponía del algoritmo CNG. Se utilizaron los algoritmos VAD y DTX del anexo B/G.729 [5]. Se obtuvieron ficheros de rastreo que contenían decisiones VAD y DTX con arreglo al procedimiento de la figura II.5 con la bandera "SYNC" activada a fin de alinear la salida con el fichero de entrada.

El códec G.711 con ruido de confort se obtuvo utilizando el procedimiento de la figura II.6. El fichero de origen fue submuestreado y ajustado en nivel por una ganancia  $G$  y a continuación codificado por la combinación del códec G.711 y el algoritmo CNG. Los datos de entrada se pusieron en memoria intermedia en tramas de 10 ms. La codificación de trama del algoritmo CNG se alineó con el comienzo del fichero vocal a fin de "sincronizarse" con el tramado correspondiente a los ficheros de rastreo VAD y DTX. Sobre la base de tramas de 10 ms, los ficheros de rastreo VAD y DTX se utilizaron para controlar el funcionamiento del algoritmo CNG. Para tramas vocales activa, se utilizó G.711 para procesar la trama de datos de entrada. Para tramas inactivas, se empleó el algoritmo CNG. La bandera DTX controló la actualización de los parámetros CNG. En el decodificador, la bandera VAD se utilizó para indicar si la trama en curso era señal vocal activa o inactiva. Se aplicó entonces una ganancia complementaria  $1/G$  (para producir un nivel de audición constante) y el resultado se sobremuestreó y almacenó como "fichero procesado".

Los resultados de este experimento ACR con ruido revelaron que, en todos los casos de interés, el códec G.711 con el algoritmo CNG de prueba actúa de manera equivalente al códec G.711 sin VAD/CNG. Se incluye aquí el caso de un fondo en calma y también los casos de fondo ruidoso (ruido de conversación confusa, de tráfico, de oficina y de la calle).

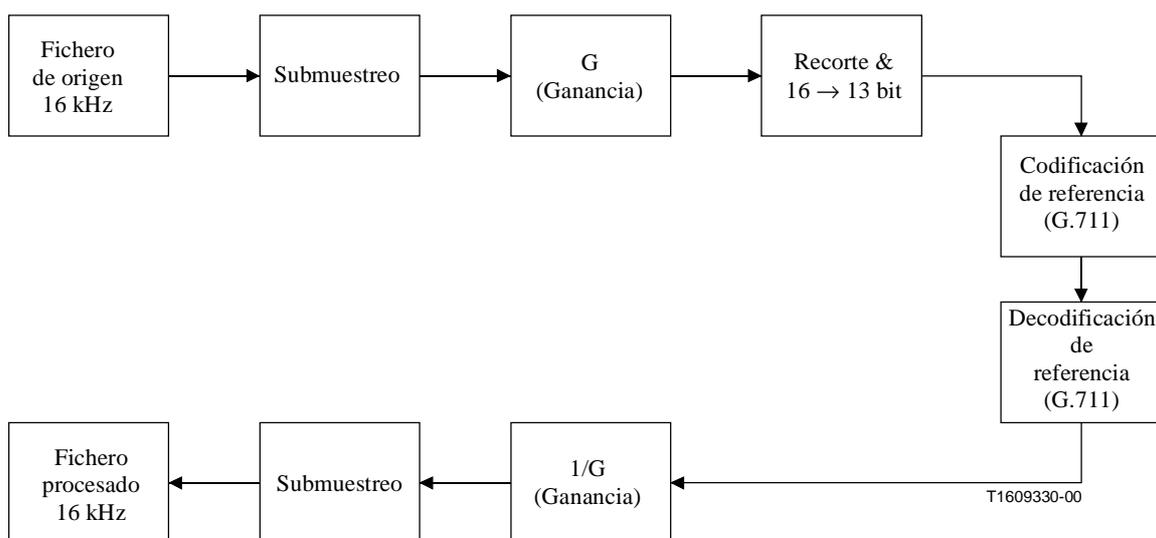
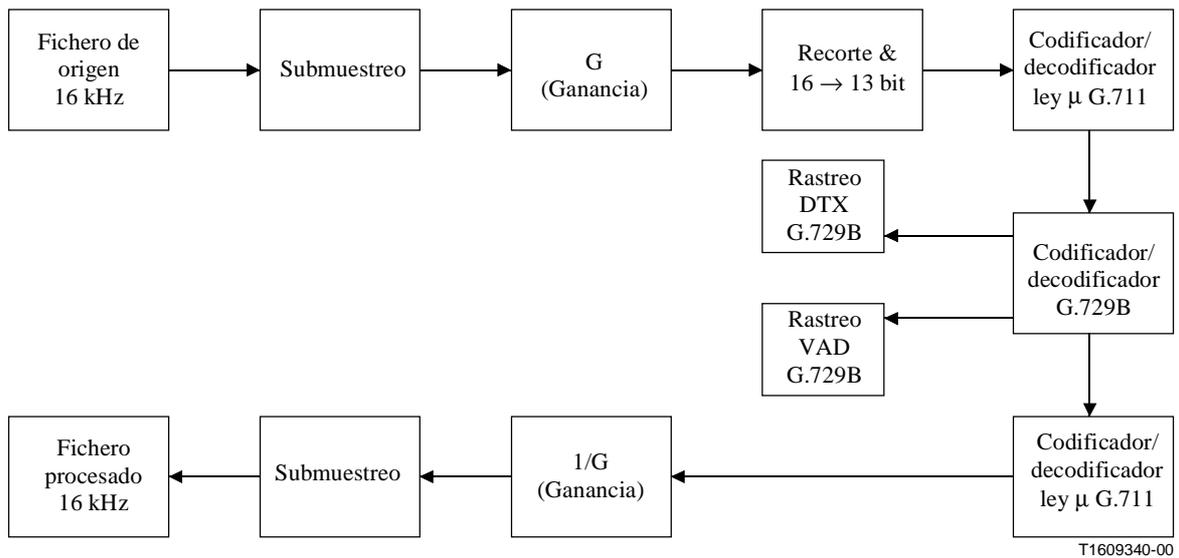
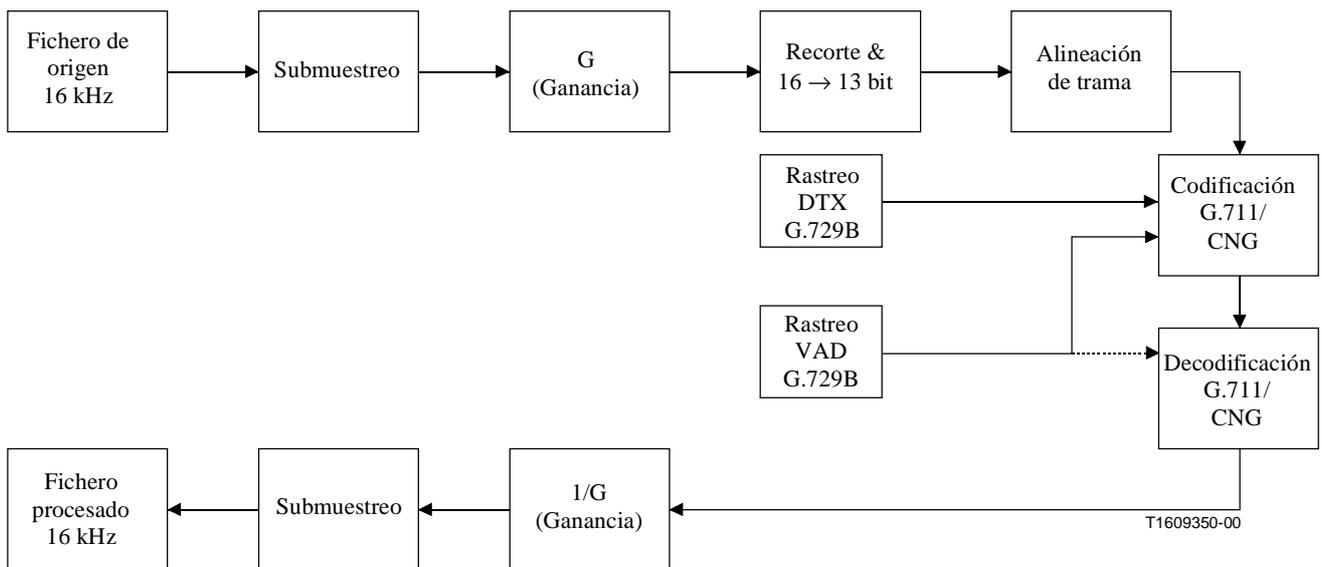


Figura II.4/G.711 – Procesamiento G.711 sin CNG



G.729B Es el anexo B a la Recomendación G.729.

**Figura II.5/G.711 – Procesamiento G.729B para obtener ficheros de rastreo VAD/DTX**



G.729B Es el anexo B a la Recomendación G.729.

**Figura II.6/G.711 – Procesamiento G.711 con CNG**

### II.5 Ejemplo de solución

Se describe en esta subcláusula un esquema de generación de ruido de confort utilizando el formato de cabida útil de ruido de confort descrito en este apéndice, que se utilizó en la evaluación descrita en II.4.

## II.5.1 Descripción del algoritmo

### II.5.1.1 Codificador

El codificador debe ser llamado cada trama por el programa llamante. Para tramas de voz activa, la señal de entrada es procesada previamente y las memorias intermedias internas se actualizan antes de volver. Para tramas inactivas, se actualizan las estimaciones de la energía de ruido de fondo y el contenido espectral. En caso de una trama SID, los parámetros estimados se cuantifican y empaquetan en la memoria intermedia de canal para su transmisión al decodificador. La velocidad de actualización del SID fue determinada por el DTX del anexo B/G.729 [5]. Los detalles del codificador CNG figuran en las subcláusulas siguientes.

#### II.5.1.1.1 Procesamiento previo

La señal de entrada es procesada previamente por un filtro IIR paso alto de primer orden para suprimir cualquier componente de baja frecuencia no deseada. El filtro paso alto viene dado por:

$$H(z) = \frac{1 - z^{-1}}{1 - (127/128)z^{-1}}$$

#### II.5.1.1.2 Análisis de autocorrelación

Los coeficientes de autocorrelación normalizados  $r_m$  y la energía de trama  $E$  se calculan sobre la base de la señal previamente procesada presentada en una ventana asimétrica de 25 ms. Para una velocidad de muestreo de 8,0 kHz, la ventana viene dada por:

$$w(n) = \begin{cases} 0,54 - 0,46 \cos\left(\frac{2\pi n}{339}\right) & n = 0, 1, \dots, 169 \\ \cos\left(\frac{2\pi(n-170)}{119}\right) & n = 170, 171, \dots, 199 \end{cases}$$

Las medias móviles de los coeficientes de autocorrelación normalizados y la energía de trama se calculan entonces para la  $i$ -ésima trama por las fórmulas:

$$\begin{aligned} \bar{r}_m(i) &= \bar{r}_m(i-1) \cdot \beta_1 + \bar{r}_m(i) \cdot (1,0 - \beta_1) & m = 1, 2, \dots, M \\ \overline{LE}(i) &= \overline{LE}(i-1) \cdot \beta_2 + \overline{LE}(i) \cdot (1,0 - \beta_2) \end{aligned}$$

donde  $LE$  es el logaritmo en base 2 de la energía de trama, y  $M$  es el orden del modelo.  $\beta_1$  y  $\beta_2$  son constantes dependientes del tamaño de trama. Si el tamaño de trama es menor o igual que 7,5 ms,  $\beta_1$  y  $\beta_2$  se fijan a 0,8, o en otro caso se fijan a 0,6. Las medias se reponen a los valores de trama vigentes si la trama anterior era de señal vocal activa.

#### II.5.1.1.3 Cálculo de los coeficientes de reflexión

El error cuadrático medio entre los coeficientes de autocorrelación instantáneos y normalizados medios se calcula por la ecuación:

$$d = \frac{1}{M} \sum_{m=1}^M (\bar{r}_m(i) - r_m(i))^2$$

Si  $d$  es menor que un umbral adaptativo  $Th$  y la última trama era inactiva, se utilizan los coeficientes promediados  $\bar{r}_m(i)$  para el cálculo de los coeficientes de reflexión; de otro modo, se utilizan los coeficientes instantáneos  $r_m(i)$ . El umbral  $Th$  se determina cada trama de acuerdo con el algoritmo siguiente.

```

if (PrevVad == 1)
    Th = 0.0
else
    Th += 0.2857 * (FRAME_SIZE / SAMPLING_RATE)
    if (Th > 0.06)
        Th = 0.06
    end
end
end

```

Los coeficientes de reflexión  $k_m(i)$  se calculan a partir de los coeficientes de autocorrelación seleccionados utilizando el algoritmo de Levinson-Durbin.

#### II.5.1.1.4 Cuantificación

Para las tramas del descriptor de inserción de silencios (SID, *silence insertion descriptor*), se cuantifican y empaquetan la energía  $\overline{LE}(i)$  y los coeficientes de reflexión  $k_m(i)$  con arreglo al formato de cabida útil especificado.

#### II.5.1.2 Decodificador

El decodificador produce ruido de confort haciendo pasar una excitación de ruido blanco escalada a través de un filtro de síntesis de predicción lineal. Los detalles siguen en las subcláusulas siguientes.

##### II.5.1.2.1 Actualización de parámetros

Los coeficientes de reflexión de la última trama SID recibida se utilizan en la trama en curso. Designemos por  $LE_{SID}$  los últimos parámetros de ruido de confort recibidos, donde la energía se ha convertido de dBoV a logaritmo en base 2. La energía utilizada en la trama en curso viene dada por:

$$LE(i) = LE(i-1) \cdot \alpha + LE_{SID} \cdot (1,0 - \alpha)$$

donde  $\alpha = 0,9$ . Este procedimiento de alisamiento se aplica para evitar cambios bruscos de la energía de la señal en el ruido de confort.

##### II.5.1.2.2 Generación de la excitación

Se utiliza un generador de número aleatorio con una distribución gaussiana para producir la secuencia  $Rn$  que es escalada por el factor  $\eta$  a la energía correcta según la ecuación:

$$\eta = \sqrt{\frac{E(i) \cdot \prod_{m=1}^M (1,0 - \hat{k}(N_m)^2)}{\frac{1}{L} \cdot \sum_{j=0}^{L-1} Rn(j)^2}}$$

donde  $L$  es la longitud de la excitación, y  $E(i)$  es la energía de trama.

Se utiliza una aproximación constante para el denominador de la ecuación citada a fin de evitar la operación producto escalar y reducir la complejidad.

### II.5.1.2.3 Síntesis de LP

Los coeficientes de reflexión se convierten en coeficientes de predicción lineal para su utilización en el filtro de síntesis de predicción lineal (LP, *linear prediction*) aplicando la recursión siguiente [6]:

$$a_i^{(i)} = -\hat{k}_i(N_i)$$

$$a_j^{(i)} = a_j^{(i-1)} + \hat{k}_i(N_i)a_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1$$

que se resuelven para  $i = 1, 2, \dots, M$ , y con el conjunto final definido como:

$$\alpha_j = a_j^{(p)} \quad 1 \leq j \leq M$$

El filtro de síntesis de predicción lineal se define como:

$$\frac{1}{A(z)} = \frac{1}{1 - \sum_{j=1}^M \alpha_j z^{-j}}$$

La excitación escalada se pasa a través del filtro para producir el ruido de confort final. La longitud de la excitación  $L$  es, en general, igual a la longitud de trama. Sin embargo, para la primera trama inactiva que sigue a una trama activa,  $L$  es igual a la longitud de trama más el orden del modelo ( $M$ ). En este caso, se ignoran las primeras muestras de salida  $M$  del filtro de síntesis.

### II.5.1.3 Retardo

No hay retardo inherente en el algoritmo de ruido de confort.

### II.5.1.4 Complejidad

El algoritmo se ha implementado en punto fijo de 16 bits utilizando la biblioteca de herramientas de soporte lógico de la UIT. La memoria y la utilización de recursos a diferentes tamaños de tramas que operan a una velocidad de muestreo de 8,0 kHz y un modelo todos polos de orden 10 se resumen en el cuadro II.2. Los WMOPS (millones ponderados de operaciones por segundo) se obtienen utilizando el contador de operaciones dentro de la biblioteca y representa el caso más desfavorable. La ROM es el tamaño estimado de un DSP de punto fijo.

**Cuadro II.2/G.711 – Necesidades de recursos CNG para un modelo de décimo orden**

| Tamaño de trama | RAM (palabras) | ROM (palabras) | WMOPS |
|-----------------|----------------|----------------|-------|
| 5 ms            | 650            | 1300           | 1,1   |
| 10 ms           | 690            | 1300           | 0,66  |
| 20 ms           | 760            | 1300           | 0,47  |

## II.5.2 Configuración probada

El algoritmo probado se especifica en el cuadro II.3.

**Cuadro II.3/G.711 – Configuración probada de CNG**

| Parámetro             | Probado |
|-----------------------|---------|
| Velocidad de muestreo | 8,0 kHz |
| Tamaño de trama       | 10 ms   |
| Orden del modelo      | 10      |
| Retardo de indagación | 5 ms    |

Se añadió una indagación de 5 ms retardando la entrada al códec vocal acompañante (G.711) como se aprecia en la figura II.7. La indagación se introdujo para ajustar adecuadamente la utilización del algoritmo VAD del anexo B/G.729 al ejemplo de solución CNG. El retardo de indagación puede evitarse en la práctica añadiendo un tiempo de retención extra al algoritmo VAD del anexo B/G.729.



**Figura II.7/G.711 – Indagación CNG durante la prueba**

## Referencias

- [1] Recomendación CCITT G.726 (1990), *Modulación por impulsos codificados diferencial adaptativa (MICDA) a 40, 32, 24, 16 kbit/s.*
- [2] Recomendación CCITT G.727 (1990), *Modulación por impulsos codificados diferencial adaptativa (MICDA) jerarquizada con 5, 4, 3 y 2 bits/muestra.*
- [3] Recomendación CCITT G.728 (1992), *Codificación de señales vocales a 16 kbits/s utilizando predicción lineal con excitación por código de bajo retardo.*
- [4] Recomendación CCITT G.722 (1988), *Codificación de audio de 7 kHz dentro de 64 kbit/s.*
- [5] Recomendación UIT-T G.729 Anexo B (1996), *Esquema de compresión de silencios para la Recomendación G.729, optimizado para terminales conformes a la Recomendación V.70.*
- [6] RABINER (L.R.), SCHAFFER (R.W.): *Digital processing of speech signals*, Prentice-Hall, 1978.
- [7] Recomendación UIT-T G.191 (1996), *Herramientas de soporte lógico para la normalización de la codificación de señales vocales y de audio.*



## SERIES DE RECOMENDACIONES DEL UIT-T

|                |   |
|----------------|---|
| Serie A        | Organización del trabajo del UIT-T  |
| Serie B        | Medios de expresión: definiciones, símbolos, clasificación  |
| Serie C        | Estadísticas generales de telecomunicaciones  |
| Serie D        | Principios generales de tarificación  |
| Serie E        | Explotación general de la red, servicio telefónico, explotación del servicio y factores humanos   |
| Serie F        | Servicios de telecomunicación no telefónicos  |
| <b>Serie G</b> | <b>Sistemas y medios de transmisión, sistemas y redes digitales</b>   |
| Serie H        | Sistemas audiovisuales y multimedios  |
| Serie I        | Red digital de servicios integrados   |
| Serie J        | Transmisiones de señales radiofónicas, de televisión y de otras señales multimedios   |
| Serie K        | Protección contra las interferencias  |
| Serie L        | Construcción, instalación y protección de los cables y otros elementos de planta exterior   |
| Serie M        | RGT y mantenimiento de redes: sistemas de transmisión, circuitos telefónicos, telegrafía, facsímil y circuitos arrendados internacionales |
| Serie N        | Mantenimiento: circuitos internacionales para transmisiones radiofónicas y de televisión  |
| Serie O        | Especificaciones de los aparatos de medida  |
| Serie P        | Calidad de transmisión telefónica, instalaciones telefónicas y redes locales  |
| Serie Q        | Conmutación y señalización  |
| Serie R        | Transmisión telegráfica   |
| Serie S        | Equipos terminales para servicios de telegrafía   |
| Serie T        | Terminales para servicios de telemática   |
| Serie U        | Conmutación telegráfica   |
| Serie V        | Comunicación de datos por la red telefónica   |
| Serie X        | Redes de datos y comunicación entre sistemas abiertos   |
| Serie Y        | Infraestructura mundial de la información y aspectos protocolo Internet   |
| Serie Z        | Lenguajes y aspectos generales de soporte lógico para sistemas de telecomunicación  |